Universal Probability Distribution for the Wave Function of an Open Quantum System

Sheldon Goldstein,[*] Joel L. Lebowitz,[†] Christian Mastrodonato,[‡] Roderich Tumulka,[§] and Nino Zanghì[¶]

April 28, 2011

Abstract

An open quantum system (i.e., one that interacts with its environment) is almost always entangled with its environment; it is therefore usually not attributed a wave function but only a reduced density matrix ρ. Nevertheless, there is a precise way of attributing to it a wave function ψ_1, called its conditional wave function, which is a random wave function of the system whose probability distribution μ_1 depends on the entangled wave function $\psi \in \mathscr{H}_1 \otimes \mathscr{H}_2$ in the Hilbert space of system and environment together. We prove several universality (or typicality) results about μ_1; they show that if the environment is sufficiently large then μ_1 does not depend much on the details of ψ and is approximately given by one of the so-called GAP measures. Specifically, for most entangled states ψ with given reduced density matrix ρ_1, μ_1 is close to $GAP(\rho_1)$. We also show that, if the coupling between the system and the environment is weak, then for most entangled states ψ from a microcanonical subspace corresponding to energies in a narrow interval $[E, E + \delta E]$ (and most bases of $\mathscr{H}_2$), μ_1 is close to $GAP(\rho_\beta)$ with ρ_β the canonical density matrix on $\mathscr{H}_1$ at inverse temperature $\beta = \beta(E)$. This provides the mathematical justification of the claim that $GAP(\rho_\beta)$ is the thermal equilibrium distribution of ψ_1.

Key words: Gaussian measures, GAP measures, Haar measure on the unitary group, thermodynamic limit, canonical ensemble in quantum mechanics, typicality theorems, conditional wave function, typical wave function.

[*]Departments of Mathematics and Physics, Rutgers University, Hill Center, 110 Frelinghuysen Road, Piscataway, NJ 08854-8019, USA. E-mail: oldstein@math.rutgers.edu

[†]Departments of Mathematics and Physics, Rutgers University, Hill Center, 110 Frelinghuysen Road, Piscataway, NJ 08854-8019, USA. E-mail: lebowitz@math.rutgers.edu

[‡]Dipartimento di Fisica, Università di Genova and INFN sezione di Genova, Via Dodecaneso 33, 16146 Genova, Italy. E-mail: christian.mastrodonato@ge.infn.it

[§]Department of Mathematics, Rutgers University, Hill Center, 110 Frelinghuysen Road, Piscataway, NJ 08854-8019, USA. E-mail: tumulka@math.rutgers.edu

[¶]Dipartimento di Fisica, Università di Genova and INFN sezione di Genova, Via Dodecaneso 33, 16146 Genova, Italy. E-mail: zanghi@ge.infn.it

Contents

1 Introduction

In this paper we establish the *universality* of certain probability distributions on Hilbert spaces known as *GAP measures* [8, 19]. This makes precise some statements and mathematical considerations outlined in our earlier paper [8] on the thermal equilibrium distribution of the wave function of an open quantum system.

By saying that GAP measures are universal we mean that the distributions μ_1 (described below) are typically close to a GAP measure,

$$\mu_1 \approx GAP \tag{1}$$

when the system's environment is sufficiently large. To illustrate the terminology of universality, one can say that the central limit theorem conveys a sense in which the Gaussian probability distribution on the real line is universal: many physically relevant probability distributions are approximately Gaussian. Instead of universality, one also often speaks of typicality; we use these two terms more or less interchangeably.

The family of GAP measures is a family of probability measures on Hilbert spaces. There is one GAP measure for every density matrix ρ in a Hilbert space $\mathscr{H}$, denoted $GAP(\rho)$; it is concentrated on the unit sphere in $\mathscr{H}$,

$$\mathbb{S}(\mathscr{H}) = \{\psi \in \mathscr{H} : \|\psi\| = 1\}\,. \tag{2}$$

The density matrix of $GAP(\rho)$ is ρ, in the usual sense that, for any probability measure μ on $\mathbb{S}(\mathscr{H})$, its density matrix is

$$\rho_\mu = \int_{\mathbb{S}(\mathscr{H})} \mu(d\psi)|\psi\rangle\langle\psi| \,, \tag{3}$$

which is also the covariance matrix of μ provided μ has mean zero. The GAP measures relevant to thermal equilibrium are those associated with canonical density matrices

$$\rho_\beta = \frac{1}{Z} e^{-\beta H} \,, \tag{4}$$

where $Z = \operatorname{tr} e^{-\beta H}$ is the normalization constant, β the inverse temperature and H the Hamiltonian. Detailed discussions of GAP measures and their physical applications can be found in [8, 19]. See [22] for a study about the support of GAP measures, that is, about what $GAP(\rho_\beta)$-distributed wave functions typically look like.

The main application of GAP measures is the characterization of the wave functions of systems we encounter in nature. In most cases we do not know a system's wave function, for example because it is a photon coming from the sun (or another star, or the cosmic microwave background, or a lamp), or because it is an electron that has escaped from a piece of metal. But in many cases the system is more or less in thermal equilibrium, and then, according to the considerations presented in [8] and here, its wave function should be GAP distributed.

1.1 Conditional Wave Function

Consider a composite quantum system consisting of two subsystems, system 1 and system 2, with associated Hilbert spaces $\mathscr{H}_1$ and $\mathscr{H}_2$. Suppose that the system is in a pure state $\psi \in \mathscr{H}_{\mathrm{total}} = \mathscr{H}_1 \otimes \mathscr{H}_2$. We ask what might be meant by the wave function of system 1. An answer is provided by the notion of *conditional wave function*, defined as follows [8]:[1] Let $b = \{b_j\}$ be an orthonormal basis of $\mathscr{H}_2$. For each choice of j, the partial inner product $\langle b_j|\psi\rangle$, taken in $\mathscr{H}_2$, is a vector belonging to $\mathscr{H}_1$. Regarding j as random (and therefore writing J), we are led to consider the random vector $\psi_1 \in \mathscr{H}_1$ given by

$$\psi_1 = \frac{\langle b_J|\psi\rangle}{\big\|\langle b_J|\psi\rangle\big\|} \tag{6}$$

where b_J is a random element of the basis $\{b_j\}$, chosen with the quantum distribution

$$\mathbb{P}^{\psi,b}(J=j) = \big\|\langle b_j|\psi\rangle\big\|^2. \tag{7}$$

[1] This definition is inspired by Bohmian mechanics, a formulation of quantum mechanics with particle trajectories, where the (non-normalized) conditional wave function ψ_1 of system 1 is defined as [4]

$$\psi_1(x) = \psi(x, Y) \tag{5}$$

for x in the configuration space of system 1, with Y the actual configuration of system 2.

We refer to ψ_1 as the conditional wave function of system 1.[2]

The distribution of ψ_1 corresponding to (6) and (7) is given by the following probability measure on $\mathbb{S}(\mathscr{H}_1)$: The probability that $\psi_1 \in A \subseteq \mathbb{S}(\mathscr{H}_1)$ is

$$\mu_1(A) = \mu_1^{\psi,b}(A) = \mathbb{P}(\psi_1 \in A) = \sum_j \big\| \langle b_j|\psi\rangle \big\|^2 \, \delta_{\langle b_j|\psi\rangle / \|\langle b_j|\psi\rangle\|}(A) \tag{9}$$

$$= \sum_j \big\| \langle b_j|\psi\rangle \big\|^2 \, 1_A\left(\frac{\langle b_j|\psi\rangle}{\|\langle b_j|\psi\rangle\|} \right), \tag{10}$$

where δ_ϕ denotes the Dirac "delta" measure (a point mass) concentrated at ϕ and 1_A denotes the characteristic function of the set A. While the density matrix ρ_{μ_1} associated with μ_1 always equals the reduced density matrix ρ_1^ψ of system 1, given by

$$\rho_1^\psi = \mathrm{tr}_2 \, |\psi\rangle\langle\psi| = \sum_j \langle b_j|\psi\rangle \langle\psi|b_j\rangle \, , \tag{11}$$

the measure μ_1 itself usually depends on the choice of the basis b, so $\mu_1 = \mu_1^{\psi,b}$.

1.2 Summary of Results

In this paper, we prove several *universality theorems* about GAP measures, Theorems 1–4, formulated in Section 2. These are statements to the effect that for *most* wave functions ψ from relevant subsets of $\mathscr{H}_1 \otimes \mathscr{H}_2$ and/or *most* orthonormal bases b of $\mathscr{H}_2$, ψ_1 is approximately GAP-distributed. Here, "most" means that the set of exceptions is small with respect to the appropriate natural uniform measure.

The basic universality property is expressed in Theorem 1, which asserts that for sufficiently large $\dim \mathscr{H}_2$, for any orthonormal basis b of $\mathscr{H}_2$, and for any density matrix ρ_1 on $\mathscr{H}_1$, most ψ in $\mathbb{S}(\mathscr{H}_1 \otimes \mathscr{H}_2)$ with the reduced density matrix $\mathrm{tr}_2 \, |\psi\rangle\langle\psi| = \rho_1$ are such that the distribution $\mu_1^{\psi,b}$ of ψ_1 is arbitrarily close to $GAP(\rho_1)$,

$$\mu_1^{\psi,b} \approx GAP(\rho_1) \, . \tag{12}$$

This fact was derived (but not rigorously proven) in Section 5.1.3 of [8]. The rigorous proof of Theorem 1 is based on the fact, found independently by several authors [24,

[2] The conditional wave function can be regarded as a precise version of the "collapsed" wave function in the standard quantum formalism: Suppose that system 1 has interacted with system 2, and their joint wave function, as produced by the appropriate Schrödinger evolution, is now

$$\sum_j c_j \psi_j^{(1)} \otimes \psi_j^{(2)} \, , \tag{8}$$

where the c_j are complex coefficients and all ψs are normalized. If system 2 is a macroscopic system and the $\psi_j^{(2)}$s are macroscopically different states then in the standard formalism one regards j as random with distribution $|c_j|^2$, and says accordingly that system 1 can be attributed the "collapsed" wave function $\psi_j^{(1)}$ with probability $|c_j|^2$. The conditional wave function of system 1, according to the above definition in the case that the $\psi_j^{(2)}$s are among the $\{b_j\}$, is indeed $\psi_j^{(1)}$ with probability $|c_j|^2$.

23, 3, 2, 15] (see also [14, 25, 10, 13]), that for a random $n \times n$ unitary matrix with distribution given by the Haar measure on the unitary group $U(n)$, the upper left (or any other) $k \times k$ submatrix, multiplied by a normalization factor $\sqrt{n}$, converges as $n \to \infty$ to a matrix of independent complex Gaussian random variables with mean 0 and variance 1. (To understand the factor $\sqrt{n}$, note that a column of a unitary $n \times n$ matrix is a unit vector, and thus a single entry should be of order $1/\sqrt{n}$.)

Theorem 2 asserts that the conclusion of Theorem 1—that (12) holds with arbitrary accuracy for sufficiently large $\dim \mathscr{H}_2$—is also true for *every* ψ with $\mathrm{tr}_2 |\psi\rangle\langle\psi| = \rho_1$ for *most* b (instead of for *every* b for *most* ψ).

Theorems 3 and 4 justify the physical conclusion that, if a system (system 1) is weakly coupled to a very large (but finite) second system (the "heat bath," system 2) then, for most wave functions of the composite system with energy in a given narrow energy range $[E, E + \delta E]$, the conditional wave function of the system is approximately GAP-distributed for most orthonormal bases of the heat bath. In more detail, let the interaction between the two systems be negligible so that the Hamiltonian can be taken to be

$$H = H_1 \otimes I_2 + I_1 \otimes H_2 \tag{13}$$

(with $I_{1/2}$ the identity operator on $\mathscr{H}_{1/2}$), and let $\mathscr{H}_R \subset \mathscr{H}_1 \otimes \mathscr{H}_2$ be a micro-canonical energy shell of the composite system, i.e., the subspace spanned by the eigenstates of the total energy with eigenvalues in $[E, E + \delta E]$. Assume that the eigenvalues of H_2 are sufficiently dense and that the dimensions of $\mathscr{H}_2$ and $\mathscr{H}_R$ are sufficiently large. Then, for most $\psi \in \mathbb{S}(\mathscr{H}_R)$,

$$\mu_1^{\psi,b} \approx GAP(\rho_\beta) \tag{14}$$

for most bases b of $\mathscr{H}_2$; here, ρ_β is the canonical density matrix (4) and $\beta = \beta(E)$.

In Theorems 3 and 4 we relax the condition that ψ have a prescribed reduced density matrix, and exploit instead *canonical typicality*. This is the fact, found independently by several groups [5, 9, 16, 17] and anticipated long before by Schrödinger [21], that for most $\psi \in \mathbb{S}(\mathscr{H}_R)$, the reduced density matrix $\mathrm{tr}_2 |\psi\rangle\langle\psi|$ is approximately of the canonical form (4). More generally, in Theorems 3 and 4 we may regard $\mathscr{H}_R$ as *any* subspace of $\mathscr{H}_1 \otimes \mathscr{H}_2$ of sufficiently high dimension. Canonical typicality then refers to the fact that for most $\psi \in \mathbb{S}(\mathscr{H}_R)$, $\mathrm{tr}_2 |\psi\rangle\langle\psi|$ is close to $\mathrm{tr}_2 \rho_R$, where ρ_R denotes $1/\dim \mathscr{H}_R$ times the projection to $\mathscr{H}_R$; the precise version of canonical typicality that we use in the proof of our Theorems 3 and 4 is due to Popescu, Short, and Winter [16, 17]. Theorem 3 asserts that for most $\psi \in \mathbb{S}(\mathscr{H}_R)$,

$$\mu_1^{\psi,b} \approx GAP(\rho_R^{(1)}) \,, \tag{15}$$

with $\rho_R^{(1)} = \mathrm{tr}_2 \rho_R$, for most bases b of $\mathscr{H}_2$.

Theorem 4 is a very similar statement but differs in the detailed meaning of "$\approx$" and refers to a fixed density matrix, such as ρ_β, in place of $\rho_R^{(1)}$ in (15).

1.3 Remarks

- *Time evolution.* It may be interesting to consider how $\mu_1^{\psi,b}$ evolves with time if the wave function $\psi = \psi_t$ of systems 1 and 2 together evolves according to the

Schrödinger equation

$$i\hbar \frac{\partial \psi_t}{\partial t} = H\psi_t \,. \qquad (16)$$

In a situation in which most $\psi \in \mathbb{S}(\mathscr{H}_R)$ have $\mu_1^{\psi,b} \approx GAP(\rho_R^{(1)})$, we may expect that even for $\psi_0 \in \mathbb{S}(\mathscr{H}_R)$ with $\mu_1^{\psi_0,b}$ far from any GAP measure, $\mu_1(t) = \mu_1^{\psi_t,b}$ will approach $GAP(\rho_R^{(1)})$ and stay near $GAP(\rho_R^{(1)})$ most of the time (though not forever, as follows from the recurrence property (almost-periodicity) of the Schrödinger evolution in a finite-dimensional Hilbert space). We leave this problem open but briefly remark that one can already conclude by interchanging the time average and the average over ψ_0 that whenever it is true for most $\psi \in \mathbb{S}(\mathscr{H}_R)$ that $\mu_1^{\psi,b} \approx GAP(\rho_R^{(1)})$, then for most $\psi_0 \in \mathbb{S}(\mathscr{H}_R)$, $\mu_1^{\psi_t,b} \approx GAP(\rho_R^{(1)})$ for most times t; the open problem is to prove a statement that concerns *all*, rather than *most*, ψ_0.

- *The role of interaction.* Another remark concerns the role of interaction (between the system and the heat bath) for obtaining the distribution $GAP(\rho_\beta)$. The nature of the interaction is relevant to our discussion in two places—although our theorems do not depend on it, as they do not mention the Hamiltonian at all. First, interaction is relevant for creating typical wave functions, as it helps evolve atypical wave functions into typical ones. This is closely related to the fact that a system coupled to a heat bath (i.e., a big second system) will typically go from non-equilibrium to thermal equilibrium only in the presence of interaction; see Section 4 of [7] for further discussion and examples. Second, it depends on the interaction which subspace of $\mathscr{H}_1 \otimes \mathscr{H}_2$ is the micro-canonical energy shell that we want $\mathscr{H}_R$ to be, and thus also which density matrix $\mathrm{tr}_2\, \rho_R$ is. In the limit of negligible interaction, $\mathrm{tr}_2\, \rho_R$ has the canonical form $\rho_\beta = (1/Z)e^{-\beta H}$, while interaction makes it deviate from this form. As a consequence of these two roles, when we want to obtain from non-equilibrium a wave function $\psi \in \mathscr{H}_1 \otimes \mathscr{H}_2$ such that the distribution of the conditional wave function ψ_1 is close to $GAP(\rho_\beta)$, we may want that the interaction be not too large (or else there will be deviations from ρ_β) and that the interaction be not too small (or else it may take too long, say longer than the present age of the universe, to reach thermal equilibrium).

1.4 Definition of the GAP Measure

For any density matrix ρ on $\mathscr{H}$, the measure $GAP(\rho)$ on (the Borel σ-algebra of) $\mathbb{S}(\mathscr{H})$ is built by starting from the measure $G(\rho)$, which is the Gaussian measure on $\mathscr{H}$ with mean 0 and covariance matrix ρ. In this paper, we are interested only in the case $\dim \mathscr{H} < \infty$. Then $G(\rho)$ can be explicitly defined as follows: Let S be the subspace of $\mathscr{H}$ on which ρ is supported, i.e., its positive spectral subspace, or equivalently the orthogonal complement of its kernel, or equivalently its range; let $d' = \dim S$ and ρ_+ the restriction of ρ to S; then $G(\rho)$ is the measure on $\mathscr{H}$ supported on S with the following density relative to the Lebesgue measure λ on S:

$$\frac{dG(\rho)}{d\lambda}(\psi) = \frac{1}{\pi^{d'} \det \rho_+} \exp(-\langle\psi|\rho_+^{-1}|\psi\rangle) \,. \qquad (17)$$

Equivalently, a $G(\rho)$-distributed random vector ψ is one whose coefficients $\langle\chi_i|\psi\rangle$ relative to an eigenbasis $\{\chi_i\}$ of ρ (i.e., $\rho\chi_i = p_i\chi_i$ with $0 \leq p_i \leq 1$) are independent complex Gaussian random variables with mean 0 and variances $\mathbb{E}|\langle\chi_i|\psi\rangle|^2 = p_i$; by a *complex Gaussian* random variable we mean one whose real and imaginary parts are independent real Gaussian random variables with equal variances.

Noting that

$$\int_{\mathscr{H}} G(\rho)(d\psi)\,\|\psi\|^2 = \operatorname{tr}\rho = 1\,, \tag{18}$$

we now define the adjusted Gaussian measure $GA(\rho)$ on $\mathscr{H}$ as:

$$GA(\rho)(d\psi) = \|\psi\|^2 G(\rho)(d\psi)\,. \tag{19}$$

If ψ^{GA} is a $GA(\rho)$-distributed vector, then $GAP(\rho)$ is the distribution of this vector projected on the unit sphere; that is, $GAP(\rho)$ is the distribution of

$$\psi^{GAP} = \frac{\psi^{GA}}{\|\psi^{GA}\|}\,. \tag{20}$$

Like $G(\rho)$ and unlike $GA(\rho)$, $GAP(\rho)$ has covariance matrix ρ.

More generally, one can define for any measure μ on $\mathscr{H}$ the "adjust-and-project" procedure. We denote by $A\mu$ the adjusted measure

$$A\mu(d\psi) = \|\psi\|^2\,\mu(d\psi)\,. \tag{21}$$

The projection on the unit sphere is defined as:

$$P : \mathscr{H} \setminus \{0\} \to \mathbb{S}(\mathscr{H})\,, \quad P(\psi) = \frac{\psi}{\|\psi\|}\,. \tag{22}$$

Then the adjusted-and-projected measure is $P_*(A\mu) = A\mu \circ P^{-1}$, where P_* denotes the action of P on measures, thus defining a mapping $P_* \circ A$ from the measures on $\mathscr{H}$ with $\int \mu(d\psi)\,\|\psi\|^2 = 1$ to the probability measures on $\mathbb{S}(\mathscr{H})$.

2 Results

2.1 GAP Measure From a Typical Wave Function of a Large System, Given the Reduced Density Matrix

Let $\mathscr{H}_{\text{total}} = \mathscr{H}_1 \otimes \mathscr{H}_2$, where $\mathscr{H}_1$ and $\mathscr{H}_2$ have respective dimension d_1 and d_2, with $d_1 < d_2 < \infty$. For any given density matrix ρ_1 on $\mathscr{H}_1$, let

$$\mathscr{R}(\rho_1) = \left\{\psi \in \mathbb{S}(\mathscr{H}_{\text{total}}) : \rho_1^\psi = \rho_1\right\} \tag{23}$$

be the set of all normalized wave functions in $\mathscr{H}_{\text{total}}$ with reduced density matrix $\rho_1^\psi = \rho_1$. We will see that $\mathscr{R}(\rho_1)$ is always non-empty.

Theorem 1 below concerns typical wave functions in $\mathscr{R}(\rho_1)$, i.e., typical wave functions with fixed reduced density matrix. The concept of "typical" refers to the uniform distribution u_{ρ_1} on $\mathscr{R}(\rho_1)$; an explicit definition of this distribution will be given in Section 3.1.

Before we formulate Theorem 1, we introduce some notation. First, for any Hilbert space $\mathscr{H}$, let $\mathscr{D}(\mathscr{H})$ denote the set of all density operators on $\mathscr{H}$, i.e., of all positive operators on $\mathscr{H}$ with trace 1. Second, when μ is a measure on $\mathscr{H}$ or $\mathbb{S}(\mathscr{H})$ and $f(\psi)$ is a measurable function on $\mathscr{H}$ or $\mathbb{S}(\mathscr{H})$ then we use the notation

$$\mu(f) := \int \mu(d\psi) f(\psi) \,. \tag{24}$$

Third, let $\|f\|_\infty = \sup_x |f(x)|$.

Theorem 1. *For every $0 < \varepsilon < 1$, $0 < \delta < 1$, and $d_1 \in \mathbb{N}$, there is $D_2 = D_2(\varepsilon, \delta, d_1) > 0$ such that for all $d_2 \in \mathbb{N}$ with $d_2 > D_2$, for every $\mathscr{H}_1$ and $\mathscr{H}_2$ with $\dim \mathscr{H}_{1/2} = d_{1/2}$, for every orthonormal basis $b = \{b_1, \ldots, b_{d_2}\}$ of $\mathscr{H}_2$, for every $\rho_1 \in \mathscr{D}(\mathscr{H}_1)$, and for every bounded measurable function $f : \mathbb{S}(\mathscr{H}_1) \to \mathbb{R}$,*

$$u_{\rho_1}\Big\{\psi \in \mathscr{R}(\rho_1) : \big|\mu_1^{\psi,b}(f) - GAP(\rho_1)(f)\big| < \varepsilon \,\|f\|_\infty\Big\} \geq 1 - \delta \,. \tag{25}$$

We give the proof, as well as those of Theorems 2–4, in Section 3.

It follows from Theorem 1 that, for every sequence $(\mathscr{H}_{2,n})_{n\in\mathbb{N}}$ of Hilbert spaces with $d_{2,n} = \dim \mathscr{H}_{2,n} \to \infty$ as $n \to \infty$ and every sequence $(b_n)_{n\in\mathbb{N}}$ of orthonormal bases $b_n = \{b_{1,n}, \ldots, b_{d_{2,n},n}\}$ of $\mathscr{H}_{2,n}$, for every $\rho_1 \in \mathscr{D}(\mathscr{H}_1)$, and for every bounded measurable function $f : \mathbb{S}(\mathscr{H}_1) \to \mathbb{R}$, the sequence of random variables $\mu_1^{\Psi_n,b_n}(f)$, where Ψ_n has distribution u_{ρ_1} on $\mathbb{S}(\mathscr{H}_1 \otimes \mathscr{H}_{2,n})$, converges in distribution, as $n \to \infty$, to the constant $GAP(\rho_1)(f)$, in fact uniformly in ρ_1, b_n and those f with $\|f\|_\infty \leq 1$. Because of the convergence for every f, we can say that the sequence of random measures $\mu_1^{\Psi_n,b_n}$ converges "weakly in distribution" to the fixed measure $GAP(\rho_1)$.

A few comments about notation. In [8], d_1 was called k, d_2 was called m, and the notation for the basis $\{b_1, \ldots, b_{d_2}\}$ was $\{|1\rangle, \ldots, |m\rangle\}$. For enumerating the basis, we will use the letter j, and thus write b_j; in [8], the notation was q_2 for j (subscript 2 because it refers to $\mathscr{H}_2$). For a random choice of j, we write J; the corresponding notation in [8] was Q_2.

2.2 GAP Measure From a Typical Basis of a Large System

As already explained in [8], instead of considering a typical wave function and a fixed basis one can consider a fixed wave function and a typical basis. Let $ONB(\mathscr{H}_2)$ be the set of all orthonormal bases of $\mathscr{H}_2$, and recall the notation $\rho_1^\psi = \mathrm{tr}_2 \,|\psi\rangle\langle\psi|$.

Theorem 2. *For every $0 < \varepsilon < 1$, $0 < \delta < 1$, $d_1 \in \mathbb{N}$, $d_2 \in \mathbb{N}$ with $d_2 > D_2(\varepsilon, \delta, d_1)$ as in Theorem 1, every $\mathscr{H}_1$ and $\mathscr{H}_2$ with $\dim \mathscr{H}_{1/2} = d_{1/2}$, every $\psi \in \mathbb{S}(\mathscr{H}_1 \otimes \mathscr{H}_2)$, and every bounded measurable function $f : \mathbb{S}(\mathscr{H}_1) \to \mathbb{R}$,*

$$u_{ONB}\Big\{(b_1, \ldots, b_{d_2}) \in ONB(\mathscr{H}_2) : \big|\mu_1^{\psi,b}(f) - GAP(\rho_1^\psi)(f)\big| < \varepsilon \,\|f\|_\infty\Big\} \geq 1 - \delta \,, \tag{26}$$

where u_{ONB} is the uniform probability measure on $ONB(\mathscr{H}_2)$, corresponding to the Haar measure on the unitary group $U(\mathscr{H}_2)$.

2.3 GAP Measure From a Typical Basis and a Typical Wave Function in a Large Subspace

In our main physical application, the reduced density matrix ρ_1^ψ is not fixed, although—by a fact known as canonical typicality—most of the relevant ψs have a reduced density matrix ρ_1^ψ that is close to a certain fixed density matrix, for example to the canonical density matrix $\rho_\beta = (1/Z)e^{-\beta H}$. In this section, we present two further universality theorems that are appropriate for such situations, in which the relevant set of ψs is a subspace of $\mathscr{H}_1 \otimes \mathscr{H}_2$ that will be denoted $\mathscr{H}_R$.

The physical setting to have in mind is this. A system with Hilbert space $\mathscr{H}_1$ is entangled with a large system whose Hilbert space is $\mathscr{H}_2$. The Hamiltonian H is thus defined on $\mathscr{H}_{\text{total}} = \mathscr{H}_1 \otimes \mathscr{H}_2$; suppose the total system is confined to a finite volume, so that H has pure point spectrum. Let $[E, E+\delta E]$ be a narrow energy window, located at a suitable energy E such as one corresponding to a more or less fixed energy per particle or per volume. Then the *micro-canonical energy shell* is the spectral subspace of H associated with this interval, i.e., the subspace spanned by the eigenvectors with eigenvalues between E and $E+\delta E$, and this is our subspace $\mathscr{H}_R$. The *micro-canonical density matrix* ρ_R is the density matrix associated with $\mathscr{H}_R$, i.e., $1/\dim \mathscr{H}_R$ times the projection to $\mathscr{H}_R$. Canonical typicality then asserts that for most wave functions in $\mathbb{S}(\mathscr{H}_R)$, the reduced density matrix is approximately ρ_β for an appropriate value of β.

For more general $\mathscr{H}_R$, canonical typicality means that for most $\psi \in \mathbb{S}(\mathscr{H}_R)$, the reduced density matrix ρ_1^ψ is close to $\operatorname{tr}_2 \rho_R$. The precise statement that we make use of is Theorem 1 of [16] or the "main theorem" of [17], which asserts, in a somewhat specialized and simplified form that suffices for our purposes:

Lemma 1. *Consider a Hilbert space $\mathscr{H}_1$ of dimension $d_1 \in \mathbb{N}$, another Hilbert space $\mathscr{H}_2$ of dimension $d_2 \in \mathbb{N}$ and a subspace $\mathscr{H}_R \subseteq \mathscr{H}_1 \otimes \mathscr{H}_2$ of dimension d_R. Let ρ_R be $1/d_R$ times the projection to $\mathscr{H}_R$, and u_R the uniform distribution on $\mathbb{S}(\mathscr{H}_R)$. Then for every $\eta > 0$,*

$$u_R\left\{\psi \in \mathbb{S}(\mathscr{H}_R) : \left\|\rho_1^\psi - \operatorname{tr}_2 \rho_R\right\|_{\text{tr}} \geq \eta + \frac{d_1}{\sqrt{d_R}}\right\} \leq 4\exp\Big(-\frac{d_R\eta^2}{18\pi^3}\Big)\,. \tag{27}$$

Here, the *trace norm* is defined by

$$\|M\|_{\text{tr}} = \operatorname{tr}|M| = \operatorname{tr}\sqrt{M^*M}\,. \tag{28}$$

By the *uniform distribution* u_R we mean the $(2d_R-1)$-dimensional surface area measure on $\mathbb{S}(\mathscr{H}_R)$, normalized so that $u_R(\mathbb{S}(\mathscr{H}_R)) = 1$.

Theorem 3. *For every $0 < \varepsilon < 1$, $0 < \delta < 1$, $d_1 \in \mathbb{N}$, every Hilbert space $\mathscr{H}_1$ with $\dim \mathscr{H}_1 = d_1$, and every continuous function $f : \mathbb{S}(\mathscr{H}_1) \to \mathbb{R}$, there is a number $D_R =$*

$D_R(\varepsilon,\delta,d_1,f)>0$ such that for all $d_R,d_2\in\mathbb{N}$ with $d_R>D_R$ and $d_2>D_2(\frac{\varepsilon}{2\|f\|_\infty},\delta/2,d_1)$ as in Theorem 1, and for every $\mathscr{H}_2$ and $\mathscr{H}_R\subseteq\mathscr{H}_1\otimes\mathscr{H}_2$ with $\dim\mathscr{H}_{2/R}=d_{2/R}$,

$$u_R\times u_{ONB}\Big\{(\psi,b)\in\mathbb{S}(\mathscr{H}_R)\times ONB(\mathscr{H}_2): \\ \big|\mu_1^{\psi,b}(f)-GAP(\mathrm{tr}_2\,\rho_R)(f)\big|<\varepsilon\Big\}\geq 1-\delta\,. \quad (29)$$

It follows that, for every sequence $(\mathscr{H}_{2,n})_{n\in\mathbb{N}}$ of Hilbert spaces with $d_{2,n}=\dim\mathscr{H}_{2,n}\to\infty$ as $n\to\infty$, every sequence $(\mathscr{H}_{R,n})_{n\in\mathbb{N}}$ of subspaces of $\mathscr{H}_1\otimes\mathscr{H}_{2,n}$ with $d_{R,n}=\dim\mathscr{H}_{R,n}\to\infty$ as $n\to\infty$, and every continuous function $f:\mathbb{S}(\mathscr{H}_1)\to\mathbb{R}$, the sequence of random variables

$$\mu_1^{\Psi_n,B_n}(f)-GAP(\mathrm{tr}_2\,\rho_{R,n})(f)\,, \quad (30)$$

where (Ψ_n,B_n) has distribution $u_{R,n}\times u_{ONB,n}$ on $\mathbb{S}(\mathscr{H}_{R,n})\times ONB(\mathscr{H}_{2,n})$, converges to zero in distribution as $n\to\infty$. We say that the sequence of random signed measures $\mu_1^{\Psi_n,B_n}-GAP(\mathrm{tr}_2\,\rho_{R,n})$ converges "weakly in distribution" to zero.

For $0<\gamma<1/\dim\mathscr{H}$ let $\mathscr{D}_{\geq\gamma}(\mathscr{H})$ denote the set of density matrices $\rho\in\mathscr{D}(\mathscr{H})$ whose eigenvalues are all greater than or equal to γ (so that, in particular, zero is not an eigenvalue of ρ).

Theorem 4. *For every $0<\varepsilon<1$, $0<\delta<1$, $d_1\in\mathbb{N}$, and $0<\gamma<1/d_1$, there are numbers $D'_R=D'_R(\varepsilon,\delta,d_1,\gamma)>0$ and $r'=r'(\varepsilon,d_1,\gamma)>0$ such that for all $d_R,d_2\in\mathbb{N}$ with $d_R>D'_R$ and $d_2>D_2(\varepsilon/2,\delta/2,d_1)$ as in Theorem 1, for every Hilbert space $\mathscr{H}_1$ with $\dim\mathscr{H}_1=d_1$, for every $\Omega\in\mathscr{D}_{\geq\gamma}(\mathscr{H}_1)$, for every $\mathscr{H}_2$ and $\mathscr{H}_R\subseteq\mathscr{H}_1\otimes\mathscr{H}_2$ with $\dim\mathscr{H}_{2/R}=d_{2/R}$ satisfying*

$$\big\|\mathrm{tr}_2(\rho_R)-\Omega\big\|_{\mathrm{tr}}<r'\,, \quad (31)$$

and for every bounded measurable function $f:\mathbb{S}(\mathscr{H}_1)\to\mathbb{R}$,

$$u_R\times u_{ONB}\Big\{(\psi,b)\in\mathbb{S}(\mathscr{H}_R)\times ONB(\mathscr{H}_2): \\ \big|\mu_1^{\psi,b}(f)-GAP(\Omega)(f)\big|<\varepsilon\,\|f\|_\infty\Big\}\geq 1-\delta\,. \quad (32)$$

If we want to consider just one particular density matrix Ω (of which zero is not an eigenvalue) then we can set γ equal to the smallest eigenvalue of Ω. It then follows that, for every sequence $(\mathscr{H}_{2,n})_{n\in\mathbb{N}}$ of Hilbert spaces with $d_{2,n}=\dim\mathscr{H}_{2,n}\to\infty$ as $n\to\infty$, and every sequence $(\mathscr{H}_{R,n})_{n\in\mathbb{N}}$ of subspaces of $\mathscr{H}_1\otimes\mathscr{H}_{2,n}$ with $d_{R,n}=\dim\mathscr{H}_{R,n}\to\infty$ and $\mathrm{tr}_2\,\rho_{R,n}\to\Omega$ as $n\to\infty$, the sequence of random measures $\mu_1^{\Psi_n,B_n}$ converges weakly in distribution to the fixed measure $GAP(\Omega)$. In short,

$$\mu_1^{\psi,b}\overset{u_R\times u_{ONB}}{\Longrightarrow}GAP(\Omega)\,. \quad (33)$$

Of the two theorems above, Theorem 3 is the simpler and perhaps more natural mathematical statement: it does not even mention any other density matrix than $\mathrm{tr}_2\,\rho_R$;

its structure is to ask first that ε, δ, and f be specified, which define the accuracy of the desired approximations;[3] and it applies to all subspaces $\mathscr{H}_R$ of sufficient dimension. For the physical application, though, we often want to compare μ_1^ψ to $GAP(\Omega)$ rather than $GAP(\operatorname{tr}_2 \rho_R)$, for example because Ω is the thermal density matrix $\rho_\beta = (1/Z)e^{-\beta H}$ while $\operatorname{tr}_2 \rho_R$ is something complicated; we usually do not need that the estimate applies uniformly to all spaces $\mathscr{H}_R$ of sufficient dimension, but instead consider only one fixed $\mathscr{H}_R$; and in that situation we can, in fact, obtain an estimate, the one provided by Theorem 4, that is uniform in f.

2.4 GAP Measure as the Thermal Equilibrium Distribution

Theorem 4 justifies regarding $GAP(\rho_\beta)$ as the thermal equilibrium distribution of the wave functions of the system 1 in the following way. Let $\mathscr{H}_R$ be the microcanonical subspace, i.e., the spectral subspace of H associated with the interval $[E, E + \delta E]$. It is a standard fact (e.g., [6, 12]) that when the interaction energy between system 1 and system 2 is sufficiently small, i.e., when we may set

$$H = H_1 \otimes I_2 + I_1 \otimes H_2 \tag{34}$$

on $\mathscr{H}_{\text{total}} = \mathscr{H}_1 \otimes \mathscr{H}_2$, and when system 2 is a large heat bath, i.e., when the eigenvalues of H_2 are sufficiently dense, then $\operatorname{tr}_2 \rho_R$ is approximately of the exponential form $Z^{-1} \exp(-\beta H_1)$ with $Z = \operatorname{tr} \exp(-\beta H_1)$ for suitable $\beta > 0$, i.e., is approximately the canonical density matrix ρ_β. Then by Theorem 4 in this special case of negligible interaction we have that for most wave functions $\psi \in \mathbb{S}(\mathscr{H}_R)$,

$$\mu_1^{\psi,b} \approx GAP(\rho_\beta) \tag{35}$$

for most orthonormal bases b of $\mathscr{H}_2$.

3 Proofs

3.1 Definition of u_{ρ_1}

According to the Schmidt decomposition [20], every $\psi \in \mathscr{H}_{\text{total}}$ can be written in the form

$$\psi = \sum_{i=1}^{d_1} c_i \, \tilde{\chi}_i \otimes \tilde{\phi}_i \tag{36}$$

where $\{\tilde{\chi}_i\}$ is an orthonormal basis in $\mathscr{H}_1$, $\{\tilde{\phi}_i\}$ is an orthonormal system in $\mathscr{H}_2$ (i.e., a set of orthonormal vectors that is not necessarily complete), and the c_i are coefficients

[3] How f defines a sense of accuracy becomes manifest if we consider finitely many test functions $f_1, \dots, f_\ell$, assume $d_R > \max(D_R(f_1), \dots, D_R(f_\ell))$, and then apply Theorem 3 to obtain that $\mu_1^{\psi,b}$ and $GAP(\operatorname{tr}_2 \rho_R)$ agree approximately on all linear combinations of $f_1, \dots, f_\ell$.

which can be chosen to be real and non-negative. If $\|\psi\| = 1$, the reduced density matrix of the system 1 is then

$$\rho_1^\psi = \sum_{i=1}^{d_1} c_i^2 |\tilde{\chi}_i\rangle\langle\tilde{\chi}_i| \,. \tag{37}$$

Thus, $\{\tilde{\chi}_i\}$ is an eigenbasis of ρ_1^ψ, and c_i^2 are the corresponding eigenvalues.

Now let a density matrix ρ_1 be given, let $\{\chi_i\}$ be an eigenbasis for ρ_1, and let $0 \leq p_i \leq 1$ be the corresponding eigenvalues. Then every $\psi \in \mathscr{R}(\rho_1)$ possesses a Schmidt decomposition of the form

$$\psi = \sum_{i=1}^{d_1} \sqrt{p_i}\, \chi_i \otimes \phi_i \tag{38}$$

with some orthonormal system $\{\phi_i\}$ in $\mathscr{H}_2$. Indeed, we know it has a Schmidt decomposition (36) in which $\{\tilde{\chi}_i\}$ is an eigenbasis of ρ_1, and c_i^2 are the eigenvalues. Reordering the terms in (36), we can make sure that $c_i = \sqrt{p_i}$. Any two eigenbases $\{\chi_i\}$ and $\{\tilde{\chi}_i\}$ of ρ_1 are related by a block unitary; more precisely, for every eigenvalue p of ρ_1, $\{\chi_i : i \in \mathscr{I}(p)\}$ and $\{\tilde{\chi}_i : i \in \mathscr{I}(p)\}$ (using the index set $\mathscr{I}(p) = \{i : c_i^2 = p\} = \{i : p_i = p\}$) are two orthonormal bases of the eigenspace of p, and thus related by a unitary matrix $(U_{ij}^{(p)})_{i,j\in\mathscr{I}(p)}$:

$$\tilde{\chi}_i = \sum_{j\in\mathscr{I}(p)} U_{ij}^{(p)}\, \chi_j \,. \tag{39}$$

Setting

$$\phi_i = \sum_{j\in\mathscr{I}(p)} U_{ji}^{(p)}\, \tilde{\phi}_j \,, \tag{40}$$

we obtain (38), and that $\{\phi_i\}$ is an orthonormal system.

Conversely, every orthonormal system $\{\phi_i\}$ in $\mathscr{H}_2$ defines, by (38), a $\psi \in \mathscr{R}(\rho_1)$. Thus, (38) defines a bijection $F_{\rho_1,\{\chi_i\}} : ONS(\mathscr{H}_2, d_1) \to \mathscr{R}(\rho_1)$. The Haar measure on the unitary group of $\mathscr{H}_2$ defines the uniform distribution on the set of orthonormal bases of $\mathscr{H}_2$, of which the uniform distribution on $ONS(\mathscr{H}_2, d_1)$ is a marginal; let $u_{\rho_1,\{\chi_i\}}$ be its image under $F_{\rho_1,\{\chi_i\}}$.

We note that $u_{\rho_1,\{\chi_i\}}$ actually does not depend on the choice of the eigenbasis $\{\chi_i\}$. Indeed, if $\{\tilde{\chi}_i\}$ is any other eigenbasis of ρ_1 (without loss of generality numbered in such a way that the eigenvalue of $\tilde{\chi}_i$ is p_i) then, as explained above, it is related to $\{\chi_i\}$ by a block unitary $d_1 \times d_1$ matrix U consisting of the blocks $(U_{ij}^{(p)})$. Let $\overline{U}$ be the matrix whose entries are the complex conjugates of the entries of U, and let $\hat{\overline{U}}$ denote the action of $\overline{U}$ on $ONS(\mathscr{H}_2, d_1)$ given by

$$\hat{\overline{U}}\Big(\big\{\phi_i : i = 1, \ldots, d_1\big\}\Big) = \bigg\{\sum_{j=1}^{d_1} \overline{U}_{ij}\phi_j : i = 1, \ldots, d_1\bigg\}. \tag{41}$$

Then

$$F_{\rho_1,\{\chi_i\}} = F_{\rho_1,\{\tilde{\chi}_i\}} \circ \hat{\hat{U}}\,. \tag{42}$$

Since the Haar measure is invariant under left multiplication, its marginal on $ONS(\mathscr{H}_2, d_1)$ is invariant under $\hat{\hat{U}}$. We thus define u_{ρ_1} to be $u_{\rho_1,\{\chi_i\}}$ for any eigenbasis $\{\chi_i\}$.

3.2 Proof of Theorem 1

We will obtain Theorem 1 from the corresponding statement about the Gaussian measures $G(\rho_1)$:

Lemma 2. *For every $0 < \tilde{\varepsilon} < 1$, $0 < \tilde{\delta} < 1$, and $d_1 \in \mathbb{N}$, there is $\tilde{D}_2 = \tilde{D}_2(\tilde{\varepsilon}, \tilde{\delta}, d_1) > 0$ such that for all $d_2 \in \mathbb{N}$ with $d_2 > \tilde{D}_2$, for every $\mathscr{H}_1$ and $\mathscr{H}_2$ with $\dim \mathscr{H}_{1/2} = d_{1/2}$, for every orthonormal basis $b = \{b_1, \ldots, b_{d_2}\}$ of $\mathscr{H}_2$, for every $\rho_1 \in \mathscr{D}(\mathscr{H}_1)$, and for every bounded measurable function $\tilde{f} : \mathscr{H}_1 \to \mathbb{R}$,*

$$u_{\rho_1}\Big\{\psi \in \mathscr{R}(\rho_1) : \big|\tilde{\mu}_1^{\psi,b}(\tilde{f}) - G(\rho_1)(\tilde{f})\big| < \tilde{\varepsilon}\,\|\tilde{f}\|_\infty\Big\} \geq 1 - \tilde{\delta}\,, \tag{43}$$

where $\tilde{\mu}_1^{\psi}$ is the distribution of $\sqrt{d_2}\langle b_J|\psi\rangle \in \mathscr{H}_1$ (not normalized) with respect to the uniform distribution of $J \in \{1, \ldots, d_2\}$.

We now collect the tools needed for the proof of Lemma 2. We note that $\tilde{\mu}_1^{\psi}$ is the sum of d_2 delta measures with equal weights,

$$\tilde{\mu}_1^{\psi} = \frac{1}{d_2}\sum_{j=1}^{d_2} \delta_{\psi_1(j)}\,, \tag{44}$$

located at the points

$$\psi_1(j) = \sqrt{d_2}\langle b_j|\psi\rangle\,. \tag{45}$$

Let $\rho_1 \in \mathscr{D}(\mathscr{H}_1)$, let $\{\chi_i\}$ be an eigenbasis of ρ_1 with eigenvalues p_i, and let $\psi \in \mathscr{R}(\rho_1)$. According to (38),

$$\psi_1(j) = \sum_{i=1}^{d_1} c_i \sqrt{d_2}\langle b_j|\phi_i\rangle \chi_i \tag{46}$$

with $c_i = \sqrt{p_i}$. Now we regard ψ as random with distribution u_{ρ_1}; then the $\psi_1(j)$ are d_2 random vectors, and $\tilde{\mu}_1^{\psi}$ is their empirical distribution. As noted above, $\{\phi_i\}$ is a random orthonormal system distributed according to a marginal of the Haar measure; in other words, the expansion coefficients $\langle b_j|\phi_i\rangle$ of

$$\phi_i = \sum_{j=1}^{d_2}\langle b_j|\phi_i\rangle b_j \tag{47}$$

form a $d_1 \times d_2$ matrix that arises as the first d_1 rows of a Haar-distributed unitary $d_2 \times d_2$ matrix.

It is therefore interesting to us how a submatrix of a Haar-distributed unitary matrix is distributed. We cite the relevant result from Olshanskij [14, Lemma 5.3] and Collins [2, Chap. 4] (the corresponding fact for orthogonal matrices was established independently by Diaconis, Eaton, and Lauritzen [3]). To formulate this result, we use the following terminology. Recall that a *complex Gaussian* random variable is one whose real and imaginary parts are independent real Gaussian random variables with equal variances. The *variation distance* of two measures μ, ν on the same σ-algebra $\mathscr{A}$ is defined to be

$$\|\mu - \nu\| = \sup_{A\in\mathscr{A}} (\mu(A) - \nu(A)) + \sup_{A\in\mathscr{A}} (\nu(A) - \mu(A)) \,. \tag{48}$$

In case μ and ν possess densities f and g relative to some measure λ on $\mathscr{A}$, this coincides with the L^1 norm of $f-g$,

$$\|\mu - \nu\| = \int \lambda(dx)\, |f(x) - g(x)| \,. \tag{49}$$

Lemma 3. *For $k, n \in \mathbb{N}$ with $k \leq n$, let the random matrix (U_{ij}) be* Haar$(U(n))$ *distributed, and let X be the upper left $k \times k$ submatrix multiplied by the normalization factor $\sqrt{n}$, $X_{ij} = \sqrt{n} U_{ij}$ for $1 \leq i, j \leq k$. Let G a random $k \times k$ matrix whose entries G_{ij} are independent complex Gaussian random variables with mean 0 and variance $\mathbb{E}|G_{ij}|^2 = 1$. Let $\mu_{k,n}$ denote the distribution of X and μ_k that of G. Then $\mu_{k,n}$ converges, as $n \to \infty$, to μ_k in the variation distance. In fact, as soon as $n \geq 2k$, $\mu_{k,n}$ and μ_k possess densities $f_{k,n}$ and f_k relative to the Lebesgue measure in $\mathbb{C}^{k\times k}$ given by*

$$f_{k,n}(X) = \mathscr{N}_{k,n}\, 1_{\|X\|_\infty < \sqrt{n}} \left(\det\Big(I - \frac{XX^*}{n} \Big) \right)^{n-2k} \tag{50}$$

(where $\mathscr{N}_{k,n}$ is the appropriate normalization factor, $\|X\|_\infty = \sup_{v\in\mathbb{C}^k\setminus\{0\}} |Xv|/|v|$, and I denotes the $k \times k$ unit matrix) and

$$f_k(X) = \frac{1}{\pi^k} e^{-\operatorname{tr} XX^*} \,, \tag{51}$$

and

$$\|f_{k,n} - f_k\|_{L^1(\mathbb{C}^{k\times k})} \to 0 \text{ as } n \to \infty. \tag{52}$$

A random matrix such as G is $\sqrt{k}$ times what is sometimes called a "standard non-selfadjoint Gaussian matrix." We will use only the following consequence of Lemma 3, concerning the convergence of the upper left $k \times 2$ entries of (U_{ij}) to a matrix of independent complex Gaussian random variables:

Corollary 1. *For every $0 < \varepsilon < 1$ and $k \in \mathbb{N}$ there is $n_0 = n_0(\varepsilon, k) > 0$ such that for every $n \in \mathbb{N}$ with $n > n_0$ and every bounded measurable function $g : \mathbb{C}^k \to \mathbb{R}$,*

$$\Big| \mathbb{E} g(\sqrt{n} U_{11}, \ldots, \sqrt{n} U_{k1}) - \mathbb{E} g(G_{11}, \ldots, G_{k1}) \Big| < \varepsilon\, \|g\|_\infty \tag{53}$$

and

$$\Big|\mathbb{E}[g(\sqrt{n}U_{11},\dots,\sqrt{n}U_{k1})g(\sqrt{n}U_{12},\dots,\sqrt{n}U_{k2})]- \\ \mathbb{E}[g(G_{11},\dots,G_{k1})g(G_{12},\dots,G_{k2})]\Big| < \varepsilon\, \|g\|_\infty^2\,, \quad (54)$$

where (U_{ij}) *is Haar distributed in* $U(n)$, *and* G_{ij} $(i=1,\dots,k;\ j=1,2)$ *are independent complex Gaussian random variables with mean 0 and variance* $\mathbb{E}|G_{ij}|^2 = 1$.

Proof. Choose $n_0 = n_0(\varepsilon,k) \geq 2k$ so that, for all $n > n_0$, $\|f_{k,n} - f_k\|_{L^1(\mathbb{C}^{k\times k})} < \varepsilon$. This is possible by Lemma 3. Then (53) and (54) follow. $\square$

Proof of Lemma 2. Set

$$\tilde{D}_2(\tilde{\varepsilon},\tilde{\delta},d_1) = \max\bigg(n_0\Big(\frac{\tilde{\varepsilon}}{2},d_1\Big), n_0\Big(\frac{\tilde{\delta}\tilde{\varepsilon}^2}{16},d_1\Big), \frac{32}{\tilde{\delta}\tilde{\varepsilon}^2}\bigg)\,. \quad (55)$$

Let us first introduce some abbreviations and notation. We write n for d_2 and k for d_1. Let $G_{11},\dots,G_{k1},G_{12},\dots,G_{k2}$ again be independent complex Gaussian random variables with mean 0 and variance $\mathbb{E}|G_{ij}|^2 = 1$. We use the basis $\{\chi_i\}$ to identify $\mathscr{H}_1$ with $\mathbb{C}^k$. Let $\tilde{f} : \mathbb{C}^k \to \mathbb{R}$ be a bounded measurable function, and let $c_i \geq 0$ be as before (the square roots of the eigenvalues of ρ_1). Set

$$g(z_1,\dots,z_k) = \tilde{f}(c_1 z_1,\dots,c_k z_k)\,. \quad (56)$$

Then g, too, is measurable and bounded with bound $\|g\|_\infty = \|\tilde{f}\|_\infty$. As we said above, the distribution of the $\langle b_j|\phi_i\rangle$ is the same as that of U_{ij} with $j \in \{1,\dots,n\}$, that is, the first k rows of a Haar distributed unitary $n\times n$ matrix. We thus write U_{ij} instead of $\langle b_j|\phi_i\rangle$, and obtain

$$\tilde{\mu}_1^{\psi}(\tilde{f}) =: \tilde{\mu}(\tilde{f}) = \frac{1}{n}\sum_{j=1}^{n} g(\sqrt{n}U_{1j},\dots,\sqrt{n}U_{kj})\,. \quad (57)$$

We write X for $\sqrt{n}$ times the upper left $k\times n$ submatrix of (U_{ij}), $\vec{X}_j$ for the j-th column of X, i.e.,

$$\vec{X}_j = (\sqrt{n}U_{1j},\dots,\sqrt{n}U_{kj})\,, \quad (58)$$

and $\vec{G}_i$ for $(G_{1i},\dots,G_{ki})$ $(i=1,2)$. In this notation,

$$\tilde{\mu}(\tilde{f}) = \frac{1}{n}\sum_{j=1}^{n} g(\vec{X}_j) \quad (59)$$

and

$$G(\rho_1)(\tilde{f}) = \mathbb{E}\tilde{f}(c_1 G_{11},\dots,c_k G_{k1}) = \mathbb{E}g(\vec{G}_1)\,. \quad (60)$$

The expression to be estimated is

$$\Big|\tilde{\mu}(\tilde{f}) - G(\rho_1)(\tilde{f})\Big|\,, \quad (61)$$

for which we have by the triangle inequality that

$$\left|\tilde{\mu}(\tilde{f}) - G(\rho_1)(\tilde{f})\right| \leq \left|\tilde{\mu}(\tilde{f}) - \mathbb{E}\tilde{\mu}(\tilde{f})\right| + \left|\mathbb{E}\tilde{\mu}(\tilde{f}) - G(\rho_1)(\tilde{f})\right|. \tag{62}$$

Note that the second contribution in (62) is nonrandom.

We first show that

$$\left|\mathbb{E}\tilde{\mu}(\tilde{f}) - G(\rho_1)(\tilde{f})\right| < \frac{\tilde{\varepsilon}}{2}\|\tilde{f}\|_\infty \tag{63}$$

if n is sufficiently large: By (59),

$$\mathbb{E}\tilde{\mu}(\tilde{f}) = \frac{1}{n}\sum_{j=1}^{n} \mathbb{E}g(\vec{X}_j) = \mathbb{E}g(\vec{X}_1) \tag{64}$$

because each $\vec{X}_j$ has the same distribution—because the columns of (U_{ij}) are exchangeable due to the invariance of the Haar measure. By (53) in Corollary 1, if $n > n_0(\tilde{\varepsilon}/2, k)$ then the absolute difference between (60) and (64) is less than $\frac{\tilde{\varepsilon}}{2}\|g\|_\infty$, i.e., (63) holds.

Concerning the first contribution in (62), Chebyshev's inequality (see, e.g., [1, p. 65]) asserts that

$$\mathbb{P}\Big(\left|\tilde{\mu}(\tilde{f}) - \mathbb{E}\tilde{\mu}(\tilde{f})\right| > \frac{\tilde{\varepsilon}}{2}\|\tilde{f}\|_\infty\Big) < \frac{4}{\tilde{\varepsilon}^2\|\tilde{f}\|_\infty^2}\mathrm{var}(\tilde{\mu}(f)), \tag{65}$$

where $\mathbb{P}$ is the Haar measure for U_{ij} and $\mathrm{var}(Y)$ is the variance of the random variable Y. Now Lemma 2 follows if we can show that

$$\mathrm{var}(\tilde{\mu}(\tilde{f})) \leq \frac{\tilde{\delta}\tilde{\varepsilon}^2}{4}\|\tilde{f}\|_\infty^2 \tag{66}$$

for sufficiently large n. We find that

$$\mathrm{var}(\tilde{\mu}(\tilde{f})) = \mathbb{E}\big[\tilde{\mu}(\tilde{f})^2\big] - (\mathbb{E}\tilde{\mu}(\tilde{f}))^2 = \frac{1}{n^2}\sum_{j,j'=1}^{n} \mathbb{E}[g(\vec{X}_j)g(\vec{X}_{j'})] - [\mathbb{E}g(\vec{X}_1)]^2. \tag{67}$$

Since the $\vec{X}_j$ are exchangeable, the joint distribution of $\vec{X}_j$ and $\vec{X}_{j'}$ for $j \neq j'$ is the same as the joint distribution of $\vec{X}_1$ and $\vec{X}_2$, so that all summands with $j \neq j'$ are equal

(and all summands with $j = j'$ are equal), and we can write

$$\operatorname{var}(\tilde{\mu}(\tilde{f})) = \frac{n^2-n}{n^2}\mathbb{E}\big[g(\vec{X}_1)\,g(\vec{X}_2)\big] + \frac{n}{n^2}\underbrace{\mathbb{E}\big[g(\vec{X}_1)^2\big]}_{\le\|g\|_\infty^2} - [\mathbb{E}g(\vec{X}_1)]^2 \tag{68}$$

$$\le \mathbb{E}\big[g(\vec{X}_1)\,g(\vec{X}_2)\big] - \frac{1}{n}\underbrace{\mathbb{E}\big[g(\vec{X}_1)\,g(\vec{X}_2)\big]}_{\ge-\|g\|_\infty^2} + \frac{1}{n}\|g\|_\infty^2 - [\mathbb{E}g(\vec{X}_1)]^2 \tag{69}$$

$$\le \mathbb{E}\big[g(\vec{X}_1)\,g(\vec{X}_2)\big] + \frac{2}{n}\|g\|_\infty^2 - [\mathbb{E}g(\vec{X}_1)]^2 \tag{70}$$

$$= \mathbb{E}\big[g(\vec{X}_1)\,g(\vec{X}_2)\big] \underbrace{-\mathbb{E}\big[g(\vec{G}_1)\,g(\vec{G}_2)\big] + \big[\mathbb{E}g(\vec{G}_1)\big]^2}_{=0} + \\ + \frac{2}{n}\|g\|_\infty^2 - [\mathbb{E}g(\vec{X}_1)]^2 \tag{71}$$

$$\le \Big|\mathbb{E}\big[g(\vec{X}_1)\,g(\vec{X}_2)\big] - \mathbb{E}\big[g(\vec{G}_1)\,g(\vec{G}_2)\big]\Big| + \frac{2}{n}\|g\|_\infty^2 + \\ + \big[\mathbb{E}g(\vec{G}_1)\big]^2 - [\mathbb{E}g(\vec{X}_1)]^2 \tag{72}$$

$$\le \Big|\mathbb{E}\big[g(\vec{X}_1)\,g(\vec{X}_2)\big] - \mathbb{E}\big[g(\vec{G}_1)\,g(\vec{G}_2)\big]\Big| + \frac{2}{n}\|g\|_\infty^2 + \\ + \big|\mathbb{E}g(\vec{G}_1) - \mathbb{E}g(\vec{X}_1)\big|\,\underbrace{\big|\mathbb{E}g(\vec{G}_1) + \mathbb{E}g(\vec{X}_1)\big|}_{\le 2\|g\|_\infty} \tag{73}$$

$$\le \frac{\tilde{\delta}\tilde{\varepsilon}^2}{16}\|g\|_\infty^2 + \frac{2}{n}\|g\|_\infty^2 + \frac{\tilde{\delta}\tilde{\varepsilon}^2}{16}\|g\|_\infty\, 2\|g\|_\infty \tag{74}$$

for $n > n_0(\tilde{\delta}\tilde{\varepsilon}^2/16, k)$ by Corollary 1. If in addition $n > 32/\tilde{\delta}\tilde{\varepsilon}^2$, we thus have that

$$\operatorname{var}(\tilde{\mu}(\tilde{f})) \le \tilde{\delta}\tilde{\varepsilon}^2\|g\|_\infty^2\Big(\frac{1}{16}+\frac{1}{16}+\frac{1}{8}\Big) = \frac{\tilde{\delta}\tilde{\varepsilon}^2}{4}\|g\|_\infty^2 = \frac{\tilde{\delta}\tilde{\varepsilon}^2}{4}\|\tilde{f}\|_\infty^2\,. \tag{75}$$

Thus, for $d_2 = n > \tilde{D}_2$ in (55), both (63) and (66) hold; Lemma 2 follows using (62) and (65). ◻

We now collect the tools needed for the proof of Theorem 1.

Lemma 4. $\tilde{\mu}_1^\psi\big(\|\cdot\|^2\big) = 1$.

Proof. Recall that $\tilde{\mu}_1^\psi$ is the distribution of $\sqrt{d_2}\langle b_J|\psi\rangle$ with $\mathbb{P}(J=j) = 1/d_2$. Thus,

$$\int_{\mathscr{H}_1} \tilde{\mu}_1^\psi(d\psi_1)\|\psi_1\|^2 = \mathbb{E}\left(d_2\|\langle b_J|\psi\rangle\|^2\right) = \sum_{j=1}^{d_2}\|\langle b_j|\psi\rangle\|^2 = \|\psi\|^2 = 1\,. \tag{76}$$

◻

Lemma 5. $\mu_1^\psi = P_* A\tilde{\mu}_1^\psi$.

Proof. By definition, $A\tilde{\mu}_1^\psi$ is the distribution of $\sqrt{d_2}\langle b_J|\psi\rangle$ with $\mathbb{P}(J=j) = \|\langle b_j|\psi\rangle\|^2$, and $P_* A\tilde{\mu}_1^\psi$ is the distribution of $\langle b_J|\psi\rangle/\|\langle b_J|\psi\rangle\|$ with $\mathbb{P}(J=j) = \|\langle b_j|\psi\rangle\|^2$. The latter is the definition of μ_1^ψ. $\square$

Recall from (18) that, for every $\rho \in \mathscr{D}(\mathscr{H})$,

$$\int_{\mathscr{H}} G(\rho)(d\psi)\,\|\psi\|^2 = 1. \tag{77}$$

Lemma 6. *For every $0 < \varepsilon < 1$ and $d \in \mathbb{N}$, there is $R = R(\varepsilon, d) > 0$ such that, for every $\mathscr{H}$ with $\dim \mathscr{H} = d$ and every $\rho \in \mathscr{D}(\mathscr{H})$,*

$$\int_{\{\psi\in\mathscr{H}:\|\psi\|<R\}} G(\rho)(d\psi)\,\|\psi\|^2 > 1-\varepsilon\,. \tag{78}$$

Proof. Let $X_1, \ldots, X_d$ be independent complex Gaussian random variables with mean 0 and variance 1. Then $X = (X_1, \ldots, X_d)$ has Gaussian distribution $G(I)$ with covariance matrix I, the identity matrix; note that

$$\int_{\mathscr{H}} G(I)(d\psi)\,\|\psi\|^2 = \mathbb{E}\sum_{i=1}^d |X_i|^2 = d. \tag{79}$$

Thus, there is $R > 0$ with

$$\int_{\{\psi\in\mathscr{H}:\|\psi\|<R\}} G(I)(d\psi)\,\|\psi\|^2 > d-\varepsilon\,. \tag{80}$$

For any $\mathscr{H}$ with $\dim \mathscr{H} = d$ and $\rho \in \mathscr{D}(\mathscr{H})$, choose an eigenbasis of ρ to identify $\mathscr{H}$ with $\mathbb{C}^d$, so $\rho = \mathrm{diag}(p_1, \ldots, p_d)$. Set $Z_i = \sqrt{p_i}X_i$, so $Z = (Z_1, \ldots, Z_d)$ has distribution $G(\rho)$. Then

$$\begin{aligned}
\int_{\{\psi\in\mathscr{H}:\|\psi\|\geq R\}} G(\rho)(d\psi)\,\|\psi\|^2 &= \mathbb{E}\Big(1_{\sum|Z_j|^2\geq R^2}\sum_i |Z_i|^2\Big) &(81)\\
&= \mathbb{E}\Big(1_{\sum p_j|X_j|^2\geq R^2}\sum_i p_i|X_i|^2\Big) &(82)\\
&\leq \mathbb{E}\Big(1_{\sum|X_j|^2\geq R^2}\sum_i |X_i|^2\Big) &(83)\\
&= \int_{\{\psi\in\mathscr{H}:\|\psi\|\geq R\}} G(I)(d\psi)\,\|\psi\|^2 < \varepsilon\,. &(84)
\end{aligned}$$

$\square$

As an abbreviation, we write $M(\mathscr{H}_1, \mathscr{H}_2, \rho_1, b, f, \varepsilon)$ or shorter $M(f, \varepsilon)$ for the set considered in (25), and $\tilde{M}(\mathscr{H}_1, \mathscr{H}_2, \rho_1, b, \tilde{f}, \tilde{\varepsilon})$ or shorter $\tilde{M}(\tilde{f}, \tilde{\varepsilon})$ for the set considered in (43).

Lemma 7. *Fix $0 < \varepsilon < 1$, $\mathscr{H}_1$, $\mathscr{H}_2$, b, and ρ_1. Let $R = R(\varepsilon/4, d_1)$ with $R(\cdot,\cdot)$ provided by Lemma 6. Then, for every bounded measurable function $f : \mathbb{S}(\mathscr{H}_1) \to \mathbb{R}$,*

$$M(f,\varepsilon) \supseteq \tilde{M}\Big((f\circ P)1_{N<R}N^2, \frac{\varepsilon}{4R^2}\Big) \cap \tilde{M}\Big(1_{N<R}N^2, \frac{\varepsilon}{4R^2}\Big), \tag{85}$$

where P denotes the projection to the unit sphere as defined in (22) *and $N(\psi_1) = \|\psi_1\|$ on $\mathscr{H}_1$.*

Note that while $f \circ P$ is not defined at $0 \in \mathscr{H}_1$, $(f\circ P)N^2$ is—as 0.

Proof of Lemma 7. If we want to show of a particular $\psi \in \mathscr{R}(\rho_1)$ that $\psi \in M(f,\varepsilon)$ then, by Lemma 5, we need to show that

$$\big|P_* A\tilde{\mu}_1^{\psi,b}(f) - GAP(\rho_1)(f)\big| < \varepsilon \, \|f\|_\infty \,. \tag{86}$$

We note that, for any probability measure $\tilde{\mu}$ on $\mathscr{H}_1$ with $\tilde{\mu}(N^2) = 1$ (in particular, by Lemma 4, for $\tilde{\mu} = \tilde{\mu}_1^{\psi,b}$), writing $g = f \circ P$, we obtain that

$$\begin{aligned}
\big|P_* A\tilde{\mu}(f) - GAP(\rho_1)(f)\big| &= \big|\tilde{\mu}(gN^2) - G(\rho_1)(gN^2)\big| \leq &(87)\\
&\leq \Big|\tilde{\mu}\big(g1_{N<R}N^2\big) - G(\rho_1)\big(g1_{N<R}N^2\big)\Big| + &(88)\\
&\quad + \Big|\tilde{\mu}\big(g1_{N\geq R}N^2\big)\Big| + &(89)\\
&\quad + \Big|G(\rho_1)\big(g1_{N\geq R}N^2\big)\Big| \,. &(90)
\end{aligned}$$

Now suppose $\psi \in \tilde{M}\Big(g1_{N<R}N^2, \varepsilon/4R^2\Big)$, which means that $\psi \in \mathscr{R}(\rho_1)$ and

$$\Big|\tilde{\mu}_1^{\psi,b}\big(g1_{N<R}N^2\big) - G(\rho_1)\big(g1_{N<R}N^2\big)\Big| < \frac{\varepsilon}{4R^2}\,\|g1_{N<R}N^2\|_\infty \,. \tag{91}$$

Since $g1_{N<R}N^2$ is a bounded function with

$$\big\|g1_{N<R}N^2\big\|_\infty \leq R^2\,\|f\|_\infty \,, \tag{92}$$

it follows from (91) that for $\tilde{\mu} = \tilde{\mu}_1^{\psi,b}$ the term (88) is less than $\|f\|_\infty\,\varepsilon/4$.

The term (89) is less than or equal to

$$\begin{aligned}
&\|f\|_\infty\,\tilde{\mu}(1_{N\geq R}N^2) &(93)\\
&= \|f\|_\infty\big(\tilde{\mu}(N^2) - \tilde{\mu}(1_{N<R}N^2)\big) &(94)\\
&= \|f\|_\infty\big(1 - \tilde{\mu}(1_{N<R}N^2)\big) &(95)\\
&\leq \|f\|_\infty\Big(1 - G(\rho_1)(1_{N<R}N^2) + \big|\tilde{\mu}(1_{N<R}N^2) - G(\rho_1)(1_{N<R}N^2)\big|\Big)\,, &(96)
\end{aligned}$$

where we have used Lemma 4. By Lemma 6,

$$G(\rho_1)(1_{N<R}N^2) > 1 - \varepsilon/4\,, \tag{97}$$

so the term (89) is less than or equal to

$$\|f\|_\infty \Big(\varepsilon/4 + \big|\tilde\mu(1_{N<R}N^2) - G(\rho_1)(1_{N<R}N^2)\big|\Big)\,. \tag{98}$$

Now suppose that also $\psi \in \tilde M\Big(1_{N<R}N^2, \varepsilon/4R^2\Big)$, which means that $\psi \in \mathscr{R}(\rho_1)$ and

$$\big|\tilde\mu_1^{\psi,b}(1_{N<R}N^2) - G(\rho_1)(1_{N<R}N^2)\big| < \frac{\varepsilon}{4R^2}\,\|1_{N<R}N^2\|_\infty = \frac{\varepsilon}{4}\,. \tag{99}$$

It follows from (98) and (99) that for $\tilde\mu = \tilde\mu_1^{\psi,b}$, the term (89) is less than $\|f\|_\infty\,\varepsilon/2$.

The term (90) is less than or equal to

$$\|f\|_\infty\, G(\rho_1)\big(1_{N\geq R}N^2\big) \tag{100}$$

$$= \|f\|_\infty\,\big(1 - G(\rho_1)(1_{N<R}N^2)\big) \tag{101}$$

$$< \|f\|_\infty\,\frac{\varepsilon}{4}\,, \tag{102}$$

where we have used $G(\rho_1)(N^2) = 1$ and (97).

Putting together the above bounds for the terms (88), (89), and (90), we obtain that their sum is bounded by $\varepsilon\|f\|_\infty(\frac14 + \frac12 + \frac14) = \varepsilon\|f\|_\infty$, so (86) is satisfied and $\psi \in M(f,\varepsilon)$. □

Proof of Theorem 1. Suppose we are given $0 < \varepsilon < 1$, $0 < \delta < 1$, and $d_1 \in \mathbb{N}$. Set

$$D_2 = D_2(\varepsilon,\delta,d_1) = \tilde D_2\Big(\frac{\varepsilon}{4R^2}, \frac{\delta}{2}, d_1\Big) \tag{103}$$

with $\tilde D_2$ as provided by Lemma 2 and R as in Lemma 7. Now consider any $d_2 \in \mathbb{N}$ with $d_2 > D_2$, any $\mathscr{H}_1$ and $\mathscr{H}_2$ with $\dim \mathscr{H}_{1/2} = d_{1/2}$, any orthonormal basis b of $\mathscr{H}_2$, any $\rho_1 \in \mathscr{D}(\mathscr{H}_1)$, and any bounded measurable function $f : \mathbb{S}(\mathscr{H}_1) \to \mathbb{R}$. Then, by two applications of Lemma 2,

$$u_{\rho_1}\Big(\tilde M\Big(g 1_{N<R}N^2, \frac{\varepsilon}{4R^2}\Big)\Big) \geq 1 - \delta/2\,, \tag{104}$$

$$u_{\rho_1}\Big(\tilde M\Big(1_{N<R}N^2, \frac{\varepsilon}{4R^2}\Big)\Big) \geq 1 - \delta/2\,. \tag{105}$$

Therefore, u_{ρ_1} of the intersection of these sets is greater than or equal to $1-\delta$. By Lemma 7, this intersection is contained in $M(f,\varepsilon)$. Thus, $u_{\rho_1}(M(f,\varepsilon)) \geq 1-\delta$, which is what we wanted to show. □

3.3 Proof of Theorem 2

Proof. Note that for any unitary U on $\mathscr{H}_2$

$$\langle U^{-1}b_j|\psi\rangle = \langle b_j|\, I_1 \otimes U\,\psi\rangle\,. \tag{106}$$

From this fact and the fact that the Haar measure is invariant under $U \mapsto U^{-1}$ it follows that the distribution of $\mu_1^{\psi,b}$, when $\psi \in \mathscr{R}(\rho_1)$ is u_{ρ_1}-distributed and b is fixed, is the same as when b is u_{ONB}-distributed and $\psi \in \mathscr{R}(\rho_1)$ is fixed. Thus, Theorem 2 is equivalent to Theorem 1. (It also follows that the distribution of $\mu_1^{\psi,b}$, when $\psi \in \mathscr{R}(\rho_1)$ is u_{ρ_1}-distributed and b is fixed, does not depend on b.) □

3.4 Continuity of GAP

For the proofs of Theorems 3 and 4, we will exploit canonical typicality, i.e., the fact that for most $\psi \in \mathbb{S}(\mathscr{H}_R)$, the reduced density matrix ρ_1^ψ is close to $\mathrm{tr}_2\, \rho_R$. Theorems 3 and 4 then follow from Theorem 2 via suitable continuity of the mapping $\rho \mapsto GAP(\rho)$. The following two lemmas provide somewhat different statements about continuity: Recall that $\mathscr{D}_{\geq\gamma}(\mathscr{H})$ is the set of density matrices with all eigenvalues greater than or equal to γ. Lemma 9 asserts that $GAP(\rho)(f)$ depends in a *uniformly* continuous way on both ρ and f when we restrict ρ to $\mathscr{D}_\gamma(\mathscr{H})$ for arbitrarily small $\gamma > 0$; continuity is not uniform without this restriction. However, Lemma 8 asserts that for any *fixed* and *continuous* test function f, continuity is uniform in ρ without restrictions.

Lemma 8. *For every $0 < \varepsilon < 1$, every $d \in \mathbb{N}$, every Hilbert space $\mathscr{H}$ with $\dim \mathscr{H} = d$, and every continuous function $f : \mathbb{S}(\mathscr{H}) \to \mathbb{R}$ there is $r = r(\varepsilon, d, f) > 0$ such that for all $\rho, \Omega \in \mathscr{D}(\mathscr{H})$,*

$$\text{if } \|\rho - \Omega\|_{\mathrm{tr}} < r \text{ then } \big|GAP(\rho)(f) - GAP(\Omega)(f)\big| < \varepsilon\,. \tag{107}$$

While all norms on $\mathscr{D}(\mathscr{H})$ are equivalent for $\dim \mathscr{H} < \infty$, we use the trace norm $\|\cdot\|_{\mathrm{tr}}$ here because in this norm the continuity extends to $\dim \mathscr{H} = \infty$ and because it is used in Lemma 1.

To formulate the other continuity statement, let $u_{\mathbb{S}(\mathscr{H})}$ denote the normalized uniform measure on the unit sphere in $\mathscr{H}$. For any density matrix $\rho \in \mathscr{D}(\mathscr{H})$ of which zero is not an eigenvalue, $GAP(\rho)$ possesses a density relative to $u_{\mathbb{S}(\mathscr{H})}$ [8].

Lemma 9. *For every $0 < \varepsilon < 1$, every $d \in \mathbb{N}$, every Hilbert space $\mathscr{H}$ with $\dim \mathscr{H} = d$, and every $0 < \gamma < 1/d$, there is $r = r(\varepsilon, d, \gamma) > 0$ such that for all $\rho, \Omega \in \mathscr{D}_{\geq\gamma}(\mathscr{H})$,*

$$\text{if } \|\rho - \Omega\|_{\mathrm{tr}} < r \text{ then } \left\|\frac{dGAP(\rho)}{du_{\mathbb{S}(\mathscr{H})}} - \frac{dGAP(\Omega)}{du_{\mathbb{S}(\mathscr{H})}}\right\|_\infty < \varepsilon\,. \tag{108}$$

As a consequence, for such ρ and Ω,

$$\big|GAP(\rho)(f) - GAP(\Omega)(f)\big| < \varepsilon\, \|f\|_1 \tag{109}$$

for every $f \in L^1\big(\mathbb{S}(\mathscr{H}), u_{\mathbb{S}(\mathscr{H})}\big)$.

It follows in particular that for any *fixed* density matrix Ω of which zero is not an eigenvalue and any sequence (ρ_n) of density matrices with $\rho_n \to \Omega$, the density of $GAP(\rho_n)$ converges to that of $GAP(\Omega)$ in the $\|\cdot\|_\infty$ norm: Take $\gamma > 0$ to be less than the smallest eigenvalue of Ω and note that only finitely many ρ_n can lie outside $\mathscr{D}_{\geq\gamma}(\mathscr{H})$.

To see that in Lemma 9 $\mathscr{D}_{\geq\gamma}(\mathscr{H})$ cannot be replaced by $\mathscr{D}(\mathscr{H})$ (i.e., that continuity is not uniform without restrictions), note that, when 0 is an eigenvalue of Ω, $GAP(\Omega)$ does not have a density with respect to $u_{\mathbb{S}(\mathscr{H})}$, so that at such an Ω, $\rho \mapsto GAP(\rho)$ is certainly not continuous in $L^\infty\big(\mathbb{S}(\mathscr{H}), u_{\mathbb{S}(\mathscr{H})}\big)$ or in the variation distance (48).

To see that in Lemma 8 one cannot drop the assumption that f is continuous, consider an Ω that has zero as an eigenvalue and a ρ that does not. Then $GAP(\Omega)$ is

concentrated on a subspace of dimension less than d while $GAP(\rho)$ has a density on the sphere and lies *near* (rather than *in*) that subspace. Thus, for a test function f that is bounded measurable but not continuous, $GAP(\rho)(f)$ does not have to be close to $GAP(\Omega)(f)$.

As part of the proof of Lemma 8, we will need the continuity property of Gaussian measures expressed in the next lemma. When μ_n, μ are measures on a topological space X, we write $\mu_n \Rightarrow \mu$ to denote that the sequence of measures μ_n converges weakly to μ. This means that $\mu_n(f) \to \mu(f)$ for every bounded continuous function $f : X \to \mathbb{R}$ and implies that the same thing is true for every bounded measurable function $f : X \to \mathbb{R}$ such that $\mu(D(f)) = 0$, where $D(f)$ is the set of discontinuities of f.

Lemma 10. *The mapping $\rho \mapsto G(\rho)$ is continuous in the weak topology on measures: If $\rho_n \in \mathscr{D}(\mathbb{C}^d)$ for every $n \in \mathbb{N}$ and $\rho_n \to \rho$ then $G(\rho_n) \Rightarrow G(\rho)$.*

Proof. We use characteristic functions; as usual, the characteristic function $\hat{\mu} : \mathbb{R}^{2d} \to \mathbb{C}$ of a probability measure μ on $\mathbb{R}^{2d}$ is defined by

$$\hat{\mu}(k_1, \ldots, k_{2d}) = \int \mu(dx_1 \cdots dx_{2d}) \exp\Bigl(i \sum_{j=1}^{2d} k_j x_j\Bigr), \tag{110}$$

or, in our notation on $\mathscr{H} = \mathbb{C}^d$,

$$\hat{\mu}(\phi) = \int \mu(d\psi) \exp\Bigl(i\mathrm{Re}\langle\phi|\psi\rangle\Bigr), \tag{111}$$

where Re denotes the real part. We write $\mu_n = G(\rho_n)$ and $\mu = G(\rho)$; their characteristic functions are:

$$\hat{\mu}_n(\psi) = \exp\bigl(-\langle\psi|\rho_n|\psi\rangle\bigr), \quad \hat{\mu}(\psi) = \exp\bigl(-\langle\psi|\rho|\psi\rangle\bigr). \tag{112}$$

If $\rho_n \to \rho$ then $\langle\psi|\rho_n|\psi\rangle \to \langle\psi|\rho|\psi\rangle$ for every ψ and thus $\hat{\mu}_n \to \hat{\mu}$ pointwise. Since (e.g., [1]) pointwise convergence of the characteristic functions is equivalent (in finite dimension) to weak convergence of the associated measures, it follows that $G(\rho_n) \Rightarrow G(\rho)$, which is what we wanted to show. □

Proof of Lemma 8. Since $\mathscr{D}(\mathscr{H})$ is compact, uniform continuity follows from continuity. That is, it suffices to show that, assuming $\rho_n \in \mathscr{D}(\mathscr{H})$ for every $n \in \mathbb{N}$,

$$\text{if } \rho_n \to \rho \text{ then } GAP(\rho_n) \Rightarrow GAP(\rho). \tag{113}$$

This follows from Lemma 10, the continuity of the adjustment mapping A defined in (21) in Section 1.4, and the continuity of the projection $P : \mathscr{H} \setminus \{0\} \to \mathbb{S}(\mathscr{H})$. Our first step is to establish the continuity of A on the set of probability measures μ on $\mathscr{H}$ such that $\int \mu(d\psi)\, \|\psi\|^2 = 1$: If, for every $n \in \mathbb{N}$, μ_n is a probability measure on the Borel σ-algebra of $\mathscr{H}$ such that $\int \mu_n(d\psi)\, \|\psi\|^2 = 1$, then

$$\text{if } \mu_n \Rightarrow \mu \text{ and } \int \mu(d\psi)\, \|\psi\|^2 = 1 \text{ then } A\mu_n \Rightarrow A\mu. \tag{114}$$

Fix $\varepsilon > 0$ and an arbitrary non-zero, bounded, continuous function $f : \mathscr{H} \to \mathbb{R}$. As before, we use the notation $N(\psi) = \|\psi\|$. Since, by hypothesis, $\mu(N^2) = 1$, there exists $R > 0$ so large that

$$\int_{\{\psi\in\mathscr{H}:\|\psi\|<R\}} \mu(d\psi)\, \|\psi\|^2 > 1 - \frac{\varepsilon}{6\|f\|_\infty}\,. \tag{115}$$

Let the "cut-off function" $\chi_0 : [0,\infty) \to [0,1]$ be any continuous function such that $\chi_0(x) = 1$ for $x \leq R$ and $\chi_0(x) = 0$ for $x \geq 2R$; set $\chi(\psi) = \chi_0(\|\psi\|)$. Because χN^2 and $f\chi N^2$ are bounded continuous functions, and because $\mu_n \Rightarrow \mu$, we have that $\mu_n(\chi N^2) \to \mu(\chi N^2)$ and $\mu_n(f\chi N^2) \to \mu(f\chi N^2)$; that is, there is an $n_1 \in \mathbb{N}$ such that, for all $n > n_1$,

$$\big|\mu_n(\chi N^2) - \mu(\chi N^2)\big| < \frac{\varepsilon}{3\|f\|_\infty} \tag{116}$$

and

$$\big|\mu_n(f\chi N^2) - \mu(f\chi N^2)\big| < \frac{\varepsilon}{3}\,. \tag{117}$$

Thus, for all $n > n_1$, we have that

$$|A\mu_n(f) - A\mu(f)| = \big|\mu_n(fN^2) - \mu(fN^2)\big| \tag{118}$$
$$\leq \big|\mu_n(f\chi N^2) - \mu(f\chi N^2)\big| + \big|\mu_n\big(f(1-\chi)N^2\big)\big| + \big|\mu\big(f(1-\chi)N^2\big)\big| \tag{119}$$
$$< \frac{\varepsilon}{3} + \|f\|_\infty \mu_n\big((1-\chi)N^2\big) + \|f\|_\infty \mu\big((1-\chi)N^2\big) \tag{120}$$
$$= \frac{\varepsilon}{3} + \|f\|_\infty\big(1 - \mu_n(\chi N^2)\big) + \|f\|_\infty\big(1 - \mu(\chi N^2)\big) \tag{121}$$
$$\leq \frac{\varepsilon}{3} + 2\|f\|_\infty\big(1 - \mu(\chi N^2)\big) + \|f\|_\infty\big|\mu_n(\chi N^2) - \mu(\chi N^2)\big| \tag{122}$$
$$\leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon\,. \tag{123}$$

This proves (114).[4]

We are now ready to establish (113). Suppose $\rho_n \to \rho$. We have that $GAP(\rho_n) = P_*A(G(\rho_n))$ and that $(AG(\rho))\,(0) = 0$. Since $\psi \mapsto P\psi$ is continuous for $\psi \neq 0$, (113) follows from (114) and Lemma 10. This completes the proof of Lemma 8. □

Proof of Lemma 9. We first note that, for any self-adjoint $d\times d$ matrix A and $\psi \in \mathbb{S}(\mathbb{C}^d)$,

$$\Big|\langle\psi|A|\psi\rangle\Big| \leq \|A\| \leq \|A\|_{\mathrm{tr}}\,. \tag{124}$$

[4] We remark that the hypothesis $\int \mu(d\psi)\|\psi\|^2 = 1$ cannot be dropped, that is, does not follow from $\int \mu_n(d\psi)\|\psi\|^2 = 1$. An example is $\mu_n = (1-1/n)\delta_0 + (1/n)\delta_{\psi_n}$, where δ_ϕ means the Dirac delta measure at ϕ and ψ_n is any vector with $\|\psi_n\|^2 = n$; then μ_n is a probability measure with $\int \mu_n(d\psi)\|\psi\|^2 = 1$ but $\mu_n \Rightarrow \delta_0$, which has $\int \delta_0(d\psi)\|\psi\|^2 = 0$.

For any density matrix $\rho \in \mathscr{D}(\mathscr{H})$ of which zero is not an eigenvalue, the density of $GAP(\rho)$ relative to $u_{\mathbb{S}(\mathscr{H})}$ is given by [8]

$$\frac{dGAP(\rho)}{du_{\mathbb{S}(\mathscr{H})}}(\psi) = \frac{1}{\pi^d \det\rho} \int_0^\infty dr\, r^{2d-1}\, r^2 \exp(-r^2 \langle\psi|\rho^{-1}|\psi\rangle) = \tag{125}$$

$$= \frac{d!}{2\pi^d \det\rho} \langle\psi|\rho^{-1}|\psi\rangle^{-d-1}. \tag{126}$$

Using the last expression, we will now show that (108) holds when ρ is sufficiently close to Ω. This follows from the facts (i) that, on $\mathscr{D}_{\geq\gamma}(\mathscr{H})$, the functions $\rho \mapsto 1/\det\rho$ and $\rho \mapsto \rho^{-1}$ are uniformly continuous, (ii) that

$$\Big|\langle\psi|\rho^{-1}|\psi\rangle - \langle\psi|\Omega^{-1}|\psi\rangle\Big| \leq \|\rho^{-1} - \Omega^{-1}\|_{\mathrm{tr}} \tag{127}$$

for all $\psi \in \mathbb{S}(\mathscr{H})$, (iii) that the function $x \mapsto x^{-d-1}$ is uniformly continuous on the interval $[1,\infty)$, and (iv) that $\langle\psi|\rho^{-1}|\psi\rangle \geq 1$, $\langle\psi|\Omega^{-1}|\psi\rangle \geq 1$. This establishes the existence of $r(\varepsilon, d, \gamma) > 0$ as described in Lemma 9.

Now (109) follows from (108) according to

$$\begin{aligned} &\big|GAP(\rho)(f) - GAP(\Omega)(f)\big| \\ &= \left|\int_{\mathbb{S}(\mathscr{H})} du_{\mathbb{S}(\mathscr{H})} \left(\frac{dGAP(\rho)}{du_{\mathbb{S}(\mathscr{H})}}(\psi) - \frac{dGAP(\Omega)}{du_{\mathbb{S}(\mathscr{H})}}(\psi)\right) f(\psi)\right| &(128)\\ &\leq \int_{\mathbb{S}(\mathscr{H})} du_{\mathbb{S}(\mathscr{H})} \left|\frac{dGAP(\rho)}{du_{\mathbb{S}(\mathscr{H})}}(\psi) - \frac{dGAP(\Omega)}{du_{\mathbb{S}(\mathscr{H})}}(\psi)\right| |f(\psi)| < \varepsilon\, \|f\|_1\,. &(129) \end{aligned}$$

□

3.5 Proof of Theorem 3

Proof of Theorem 3. Suppose we are given $0 < \varepsilon < 1$, $0 < \delta < 1$, $d_1 \in \mathbb{N}$, a Hilbert space $\mathscr{H}_1$ of dimension d_1, and a continuous function $f : \mathbb{S}(\mathscr{H}_1) \to \mathbb{R}$. Set

$$D_R = D_R(\varepsilon, \delta, d_1, f) = \frac{4}{r(\varepsilon/2, d_1, f)^2} \max\Big(d_1^2, 18\pi^3 \log(8/\delta)\Big), \tag{130}$$

with $r(\varepsilon, d, f)$ as provided by Lemma 8. Now consider any $d_R, d_2 \in \mathbb{N}$ with $d_R > D_R$ and $d_2 > D_2(\varepsilon/2\|f\|_\infty, \delta/2, d_1)$, any $\mathscr{H}_2$ and $\mathscr{H}_R \subseteq \mathscr{H}_1 \otimes \mathscr{H}_2$ with $\dim \mathscr{H}_{2/R} = d_{2/R}$. Let $M(f, \varepsilon)$ be the set mentioned in (29),

$$M(f,\varepsilon) = \Big\{\big(\psi, b\big) \in \mathbb{S}(\mathscr{H}_R) \times ONB(\mathscr{H}_2) : \big|\mu_1^{\psi,b}(f) - GAP(\mathrm{tr}_2\, \rho_R)(f)\big| < \varepsilon\Big\}, \tag{131}$$

let

$$M'(f,\varepsilon) = \Big\{\big(\psi, b\big) \in \mathbb{S}(\mathscr{H}_R) \times ONB(\mathscr{H}_2) : \big|\mu_1^{\psi,b}(f) - GAP(\rho_1^\psi)(f)\big| < \varepsilon\Big\} \tag{132}$$

and

$$M''(\varepsilon) = \Big\{\psi \in \mathbb{S}(\mathscr{H}_R) : \|\rho_1^\psi - \operatorname{tr}_2 \rho_R\|_{\mathrm{tr}} < \varepsilon\Big\}. \tag{133}$$

Then, by Lemma 8,

$$M(f,\varepsilon) \supseteq M'\Big(f, \frac{\varepsilon}{2}\Big) \cap \Big[M''\Big(r\big(\frac{\varepsilon}{2}, d_1, f\big)\Big) \times ONB(\mathscr{H}_2)\Big]\,. \tag{134}$$

Theorem 2 yields, using our assumption $d_2 > D_2(\varepsilon/2\|f\|_\infty, \delta/2, d_1)$, that for every $\psi \in \mathbb{S}(\mathscr{H}_R)$,

$$u_{ONB}\Big\{b \in ONB(\mathscr{H}_2) : \big|\mu_1^{\psi,b}(f) - GAP(\rho_1^\psi)(f)\big| < \frac{\varepsilon}{2}\Big\} \geq 1 - \delta/2\,. \tag{135}$$

Thus, averaging over $\psi \in \mathbb{S}(\mathscr{H}_R)$ according to u_R,

$$u_R \times u_{ONB}\Big(M'(f,\varepsilon/2)\Big) \geq 1 - \delta/2\,. \tag{136}$$

Lemma 1 with $\eta = r/2$ for $r = r(\varepsilon/2, d_1, f)$ yields, using our assumption $d_R > 4d_1^2/r^2$, which implies that $d_1/\sqrt{d_R} \leq r/2$, that

$$u_R(M''(r)) \geq 1 - 4\exp\Big(-\frac{d_R r^2}{18\pi^3 4}\Big)\,. \tag{137}$$

Using our assumption $d_R > 18\pi^3 4 \log(8/\delta)/r^2$, the right hand side is greater than or equal to $1 - \delta/2$, and thus

$$u_R \times u_{ONB}\Big[M''(r) \times ONB(\mathscr{H}_2)\Big] \geq 1 - \delta/2\,. \tag{138}$$

From (136), (138), and (134) together we have that

$$u_R \times u_{ONB}\Big[M(f,\varepsilon)\Big] \geq 1 - \delta\,, \tag{139}$$

which is what we wanted to show. □

3.6 Proof of Theorem 4

Proof of Theorem 4. Suppose we are given $0 < \varepsilon < 1$, $0 < \delta < 1$, $d_1 \in \mathbb{N}$, $0 < \gamma < 1/d_1$, and a Hilbert space $\mathscr{H}_1$ of dimension d_1. Set

$$D'_R = D'_R(\varepsilon,\delta,d_1,\gamma) = \frac{4}{(r')^2}\max\Big(d_1^2, 18\pi^3\log(8/\delta)\Big)\,, \tag{140}$$

$$r' = r'(\varepsilon, d_1, \gamma) = \frac{1}{2} r(\varepsilon/2, d_1, \gamma)\,, \tag{141}$$

with $r(\varepsilon, d, \gamma)$ as provided by Lemma 9. Now consider any $d_R, d_2 \in \mathbb{N}$ with $d_R > D'_R$ and $d_2 > D_2(\varepsilon/2, \delta/2, d_1)$, any $\Omega \in \mathscr{D}_{\geq\gamma}(\mathscr{H}_1)$, any $\mathscr{H}_2$ and $\mathscr{H}_R \subseteq \mathscr{H}_1 \otimes \mathscr{H}_2$ with

$\dim \mathscr{H}_{2/R} = d_{2/R}$, and any bounded measurable function $f : \mathbb{S}(\mathscr{H}_1) \to \mathbb{R}$. Let $M_0(f,\varepsilon)$ be the set mentioned in (32),

$$M_0(f,\varepsilon) = \Big\{ \big(\psi, b\big) \in \mathbb{S}(\mathscr{H}_R) \times ONB(\mathscr{H}_2) : \big|\mu_1^{\psi,b}(f) - GAP(\Omega)(f)\big| < \varepsilon \, \|f\|_\infty \Big\}, \tag{142}$$

let, as in the proof of Theorem 3,

$$M'(f,\varepsilon) = \Big\{ \big(\psi, b\big) \in \mathbb{S}(\mathscr{H}_R) \times ONB(\mathscr{H}_2) : \big|\mu_1^{\psi,b}(f) - GAP(\rho_1^{\psi})(f)\big| < \varepsilon \Big\}, \tag{143}$$

let

$$M_0''(\varepsilon) = \Big\{ \psi \in \mathbb{S}(\mathscr{H}_R) : \|\rho_1^{\psi} - \Omega\|_{\mathrm{tr}} < \varepsilon \Big\}, \tag{144}$$

and let, as in the proof of Theorem 3,

$$M''(\varepsilon) = \Big\{ \psi \in \mathbb{S}(\mathscr{H}_R) : \|\rho_1^{\psi} - \mathrm{tr}_2\, \rho_R\|_{\mathrm{tr}} < \varepsilon \Big\}. \tag{145}$$

Now assume $\big\|\mathrm{tr}_2\, \rho_R - \Omega\big\|_{\mathrm{tr}} < r'$. Then

$$M_0''(2r') \supseteq M''(r') \tag{146}$$

and, by Lemma 9 and $\|f\|_1 \leq \|f\|_\infty$,

$$M_0(f,\varepsilon) \supseteq M'\bigg(f, \frac{\varepsilon \|f\|_\infty}{2}\bigg) \cap \Big[M_0''(2r') \times ONB(\mathscr{H}_2) \Big] . \tag{147}$$

As in the proof of Theorem 3, Theorem 2 yields (136) with ε replaced by $\varepsilon\|f\|_\infty$ using our assumption $d_2 > D_2(\varepsilon/2, \delta/2, d_1)$, and Lemma 1 yields (138) with r replaced by r', using our assumption $d_R > D_R'$. From (136), (138), (146), and (147) together we have that

$$u_R \times u_{ONB}\Big[M_0(f,\varepsilon) \Big] \geq 1 - \delta \,, \tag{148}$$

which is what we wanted to show. □

Acknowledgments. We are grateful to Benoit Collins for helpful discussions. S. Goldstein was supported in part by the National Science Foundation [grant DMS-0504504]. J. L. Lebowitz and C. Mastrodonato were supported in part by the National Science Foundation [grant DMR 08-02120] and the Air Force Office of Scientific Research [grant AF-FA 49620-01-0154]. N. Zanghì was supported in part by Istituto Nazionale di Fisica Nucleare.

References

[1] P. Billingsley: *Probability and Measure.* John Wiley & Sons, 1986.

[2] B. Collins: *Intégrales matricielles et Probabilités Non-Commutatives.* Ph. D. thesis, Department of Mathematics, Université Paris 6, 2003. http://tel.archives-ouvertes.fr/docs/00/04/59/88/PDF/tel-00004306.pdf

[3] P. Diaconis, M. L. Eaton, S. L. Lauritzen: Finite de Finetti Theorems in Linear Models and Multivariate Analysis. *Scand. J. Statist.* **19**: 289–315, 1992.

[4] D. Dürr, S. Goldstein, N. Zanghì: Quantum equilibrium and the origin of absolute uncertainty. *J. Statist. Phys.* **67**: 843–907, 1992. http://arxiv.org/abs/quant-ph/0308039

[5] J. Gemmer, G. Mahler, M. Michel: *Quantum Thermodynamics: Emergence of Thermodynamic Behavior within Composite Quantum Systems.* Lecture Notes in Physics **657**. Berlin: Springer, 2004.

[6] H.-O. Georgii: The equivalence of ensembles for classical systems of particles. *J. Statist. Phys.* **80**: 1341–1378, 1995.

[7] S. Goldstein, J. L. Lebowitz, C. Mastrodonato, R. Tumulka, N. Zanghì: Approach to Thermal Equilibrium of Macroscopic Quantum Systems. *Physical Review E* **81**: 011109, 2010. http://arxiv.org/abs/0911.1724

[8] S. Goldstein, J.L. Lebowitz, R. Tumulka, N. Zanghì: On the Distribution of the Wave Function for System in Thermal Equilibrium. *J. Statist. Phys.* **125**: 1193–1221, 2006. http://arxiv.org/abs/quant-ph/0309021

[9] S. Goldstein, J.L. Lebowitz, R. Tumulka, N. Zanghì: Canonical Typicality. *Phys. Rev. Lett.* **96**: 050403, 2006. http://arxiv.org/abs/cond-mat/0511091

[10] T. Jiang: How Many Entries of a Typical Orthogonal Matrix Can Be Approximated by Independent Normals? *Ann. Probab.* **34(4)**: 1497–1529, 2006. http://arxiv.org/abs/math/0601457

[11] N. Linden, S. Popescu, A. J. Short, A. Winter: Quantum mechanical evolution towards thermal equilibrium. *Phys. Rev. E* **79**: 061103, 2009. http://arxiv.org/abs/0812.2385

[12] A. Martin-Löf: *Statistical Mechanics and the Foundations of Thermodynamics* (Lecture Notes in Physics **101**). Berlin: Springer, 1979.

[13] C. Mastrodonato, R. Tumulka: Elementary Proof for Asymptotics of Large Haar-Distributed Unitary Matrices. *Lett. Math. Phys.* **82**: 51–59, 2007. http://arxiv.org/abs/0705.3146

[14] G. I. Ol'shanskij: Unitary representations of infinite dimensional pairs (G, K) and the formalism of R. Howe. Pages 269–463 in A.M. Vershik and D.P. Zhelobenko (editors), *Representation of Lie groups and related topics*, Volume 7 of the series

Advanced Studies in Contemporary Mathematics. New York: Gordon and Breach, 1990.

[15] D. Petz, J. Réffy: On asymptotics of large Haar distributed unitary matrices. *Periodica Mathematica Hungarica* **49**: 103–117, 2004. http://arxiv.org/abs/math.PR/0310338

[16] S. Popescu, A. J. Short, A. Winter: The foundations of statistical mechanics from entanglement: Individual states vs. averages. Preprint http://arxiv.org/abs/quant-ph/0511225, 2005.

[17] S. Popescu, A. J. Short, A. Winter: Entanglement and the foundation of statistical mechanics. *Nature Physics* **21(11)**: 754–758, 2006.

[18] P. Reimann: Typicality for Generalized Microcanonical Ensembles. *Phys. Rev. Lett.* **99**: 160404, 2007. http://arxiv.org/abs/0710.4214

[19] P. Reimann: Typicality of pure states randomly sampled according to the Gaussian adjusted projected measure. *J. Statist. Phys.* **132(5)**: 921–935, 2008. http://arxiv.org/abs/0805.3102

[20] Schmidt decomposition. In *Wikipedia, the free encyclopedia* (accessed December 19, 2009), http://en.wikipedia.org/wiki/Schmidt_decomposition

[21] E. Schrödinger: *Statistical Thermodynamics.* Second Edition, Cambridge: University Press, 1952.

[22] R. Tumulka, N. Zanghì: Smoothness of Wave Functions in Thermal Equilibrium. *J. Math. Phys.* **46**: 112104, 2005. http://arxiv.org/abs/math-ph/0509028

[23] D. Weingarten: Asymptotic behavior of group integrals in the limit of infinite rank. *J. Math. Phys.* **19**: 999–1001, 1978.

[24] Y. Yamasaki: Projective limit of Haar measures on $O(n)$. *Publications of the Research Institute for Mathematical Sciences at Kyoto University* **8**: 141–149, 1972.

[25] K. Zyczkowski, H.-J. Sommers: Truncations of random unitary matrices. *J. Phys. A: Math. Gen.* **33**: 2045–2057, 2000. http://arxiv.org/abs/chao-dyn/9910032